

0500

OPE #2

## RAW SEQUENCE LISTING

DATE: 07/10/2000

PATENT APPLICATION: US/09/604,231

TIME: 15:45:58

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07102000\I604231.raw

ENTERED

```

3 <110> APPLICANT: Pompejus, Markus
4 Kroger, Burkhard
5 Schroder, Hartwig
6 Zelder, Oskar
7 Haberhauer, Gregor
W--> 8 <120> TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
W--> 9 PHOSPHOENOLPYRUVATE:SUGAR PHOSPHOTRANSFERASE
W--> 10 SYSTEM PROTEINS
W--> 11 <130> FILE REFERENCE: BGI-122CP
W--> 12 <140> CURRENT APPLICATION NUMBER:
C--> 13 <141> CURRENT FILING DATE: 2000-06-27
W--> 14 <160> NUMBER OF SEQ ID: 34
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1527
18 <212> TYPE: DNA
19 <213> ORGANISM: Corynebacterium glutamicum
21 <220> FEATURE:
22 <221> NAME/KEY: CDS
23 <222> LOCATION: (101)..(1504)
24 <223> OTHER INFORMATION: RXS00315
26 <400> SEQUENCE: 1
27 ctcattgcat ctgcgcggtt cgcgttcttg ccagtgttgg ttggtttcac cgcaaccaag 60
29 cgtttcggcg gcaatgagtt cctgggggcc gcgtattggt atg gcg atg gtg ttc 115
30 Met Ala Met Val Phe
31 1 5
33 ccg agc ttg gtg aac ggc tac gac gtg gcc gcc acc atg gct gcg ggc 163
34 Pro Ser Leu Val Asn Gly Tyr Asp Val Ala Ala Thr Met Ala Ala Gly
35 10 15 20
37 gaa atg cca atg tgg tcc ctg ttt ggt tta gat gtt gcc caa gcc ggt 211
38 Glu Met Pro Met Trp Ser Leu Phe Gly Leu Asp Val Ala Gln Ala Gly
39 25 30 35
41 tac cag ggc acc gtg ctt cct gtg ctg gtg gtt tct tgg att ctg gca 259
42 Tyr Gln Gly Thr Val Leu Pro Val Leu Val Val Ser Trp Ile Leu Ala
43 40 45 50
45 acg atc gag aag ttc ctg cac aag cga ctc aag ggc act gca gac ttc 307
46 Thr Ile Glu Lys Phe Leu His Lys Arg Leu Lys Gly Thr Ala Asp Phe
47 55 60 65
49 ctg atc act cca gtg ctg acg ttg ctg ctc acc gga ttc ctt aca ttc 355
50 Leu Ile Thr Pro Val Leu Thr Leu Leu Thr Gly Phe Leu Thr Phe
51 70 75 80 85
53 atc gcc att ggc cca gca atg cgc tgg gtg ggc gat gtg ctg gca cac 403
54 Ile Ala Ile Gly Pro Ala Met Arg Trp Val Gly Asp Val Leu Ala His
55 90 95 100
57 ggt cta cag gga ctt tat gat ttc ggt ggt cca gtc ggc ggt ctg ctc 451
58 Gly Leu Gln Gly Leu Tyr Asp Phe Gly Gly Pro Val Gly Gly Leu Leu
59 105 110 115
61 ttc ggt ctg gtc tac tca cca atc gtc atc act ggt ctg cac cag tcc 499

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/604,231

DATE: 07/10/2000  
 TIME: 15:45:58

Input Set : A:\seqlist.txt  
 Output Set: N:\CRF3\07102000\I604231.raw

```

62 Phe Gly Leu Val Tyr Ser Pro Ile Val Ile Thr Gly Leu His Gln Ser
63      120      125      130
65 ttc ccg cca att gag ctg gag ttt aac cag ggt gga tcc ttc atc 547
66 Phe Pro Pro Ile Glu Leu Glu Leu Phe Asn Gln Gly Gly Ser Phe Ile
67      135      140      145
69 ttc gca acg gca tct atg gct aat atc gcc cag ggt gcg gca tgt ttg 595
70 Phe Ala Thr Ala Ser Met Ala Asn Ile Ala Gln Gly Ala Ala Cys Leu
71 150      155      160      165
73 gca gtg ttc ttc ctg gcg aag agt gaa aag ctc aag ggc ctt gca ggt 643
74 Ala Val Phe Phe Leu Ala Lys Ser Glu Lys Leu Lys Gly Leu Ala Gly
75      170      175      180
77 gct tca ggt gtc tcc gct gtt ctt ggt att acg gag cct gcg atc ttc 691
78 Ala Ser Gly Val Ser Ala Val Leu Gly Ile Thr Glu Pro Ala Ile Phe
79      185      190      195
81 ggt gtg aac ctt cgc ctg cgc tgg ccg ttc ttc atc ggt atc ggt acc 739
82 Gly Val Asn Leu Arg Leu Arg Trp Pro Phe Phe Ile Gly Ile Gly Thr
83      200      205      210
85 gca gct atc ggt ggc gct ttg att gca ctc ttt aat atc aag gca gtt 787
86 Ala Ala Ile Gly Gly Ala Leu Ile Ala Leu Phe Asn Ile Lys Ala Val
87      215      220      225
89 gcg ttg ggc gct gca ggt ttc ttg ggt gtt gtt tct att gat gct cca 835
90 Ala Leu Gly Ala Ala Gly Phe Leu Gly Val Val Ser Ile Asp Ala Pro
91 230      235      240      245
93 gat atg gtc atg ttc ttg gtg tgt gca gtt gtt acc ttc ttc atc gca 883
94 Asp Met Val Met Phe Leu Val Cys Ala Val Val Thr Phe Phe Ile Ala
95      250      255      260
97 ttc ggc gca gcg att gct tat ggc ctt tac ttg gtt cgc cgc aac ggc 931
98 Phe Gly Ala Ala Ile Ala Tyr Gly Leu Tyr Leu Val Arg Arg Asn Gly
99      265      270      275
101 agc att gat cca gat gca acc gct gct cca gtg cct gca gga acg acc 979
102 Ser Ile Asp Pro Asp Ala Thr Ala Ala Pro Val Pro Ala Gly Thr Thr
103      280      285      290
105 aaa gcc gaa gca gaa gca ccc gca gaa ttt tca aac gat tcc acc atc 1027
106 Lys Ala Glu Ala Glu Ala Pro Ala Glu Phe Ser Asn Asp Ser Thr Ile
107      295      300      305
109 atc cag gca cct ttg acc ggt gaa gct att gca ctg agc agc gtc agc 1075
110 Ile Gln Ala Pro Leu Thr Gly Glu Ala Ile Ala Leu Ser Ser Val Ser
111 310      315      320      325
113 gat gcc atg ttt gcc agc gga aag ctt ggc tcg ggc gtt gcc atc gtc 1123
114 Asp Ala Met Phe Ala Ser Gly Lys Leu Gly Ser Gly Val Ala Ile Val
115      330      335      340
117 cca acc aag ggg cag tta gtt tct ccg gtg agt gga aag att gtg gtg 1171
118 Pro Thr Lys Gly Gln Leu Val Ser Pro Val Ser Gly Lys Ile Val Val
119      345      350      355
121 gca ttc cca tct ggc cat gct ttc gca gtt cgc acc aag gct gag gat 1219
122 Ala Phe Pro Ser Gly His Ala Phe Ala Val Arg Thr Lys Ala Glu Asp
123      360      365      370
125 ggt tcc aat gtg gat atc ttg atg cac att ggt ttc gac aca gta aac 1267
126 Gly Ser Asn Val Asp Ile Leu Met His Ile Gly Phe Asp Thr Val Asn

```

## RAW SEQUENCE LISTING

DATE: 07/10/2000

PATENT APPLICATION: US/09/604,231

TIME: 15:45:58

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07102000\I604231.raw

```

127      375      380      385
129 ctc aac ggc acg cac ttt aac ccg ctg aag aag cag ggc gat gaa gtc 1315
130 Leu Asn Gly Thr His Phe Asn Pro Leu Lys Lys Gln Gly Asp Glu Val
131 390      395      400      405
133 aaa gca ggg gag ctg ctg tgt gaa ttc gat att gat gcc att aag gct 1363
134 Lys Ala Gly Glu Leu Leu Cys Glu Phe Asp Ile Asp Ala Ile Lys Ala
135      410      415      420
137 gca ggt tat gag gta acc acg ccg att gtt gtt tcg aat tac aag aaa 1411
138 Ala Gly Tyr Glu Val Thr Thr Pro Ile Val Val Ser Asn Tyr Lys Lys
139      425      430      435
141 acc gga cct gta aac act tac ggt ttg ggc gaa att gaa gcg gga gcc 1459
142 Thr Gly Pro Val Asn Thr Tyr Gly Leu Gly Glu Ile Glu Ala Gly Ala
143      440      445      450
145 aac ctg ctc aac gtc gca aag aaa gaa gcg gtg cca gca aca cca 1504
146 Asn Leu Leu Asn Val Ala Lys Lys Glu Ala Val Pro Ala Thr Pro
147      455      460      465
149 taagttgaaa cottgagtggt tcg 1527
152 <210> SEQ ID NO: 2
153 <211> LENGTH: 468
154 <212> TYPE: PRT
155 <213> ORGANISM: Corynebacterium glutamicum
157 <400> SEQUENCE: 2
158 Met Ala Met Val Phe Pro Ser Leu Val Asn Gly Tyr Asp Val Ala Ala
159 1 5 10 15
161 Thr Met Ala Ala Gly Glu Met Pro Met Trp Ser Leu Phe Gly Leu Asp
162 20 25 30
164 Val Ala Gln Ala Gly Tyr Gln Gly Thr Val Leu Pro Val Leu Val Val
165 35 40 45
167 Ser Trp Ile Leu Ala Thr Ile Glu Lys Phe Leu His Lys Arg Leu Lys
168 50 55 60
170 Gly Thr Ala Asp Phe Leu Ile Thr Pro Val Leu Thr Leu Leu Thr
171 65 70 75 80
173 Gly Phe Leu Thr Phe Ile Ala Ile Gly Pro Ala Met Arg Trp Val Gly
174 85 90 95
176 Asp Val Leu Ala His Gly Leu Gln Gly Leu Tyr Asp Phe Gly Gly Pro
177 100 105 110
179 Val Gly Gly Leu Phe Gly Leu Val Tyr Ser Pro Ile Val Ile Thr
180 115 120 125
182 Gly Leu His Gln Ser Phe Pro Pro Ile Glu Leu Glu Leu Phe Asn Gln
183 130 135 140
185 Gly Gly Ser Phe Ile Phe Ala Thr Ala Ser Met Ala Asn Ile Ala Gln
186 145 150 155 160
188 Gly Ala Ala Cys Leu Ala Val Phe Phe Leu Ala Lys Ser Glu Lys Leu
189 165 170 175
191 Lys Gly Leu Ala Gly Ala Ser Gly Val Ser Ala Val Leu Gly Ile Thr
192 180 185 190
194 Glu Pro Ala Ile Phe Gly Val Asn Leu Arg Leu Arg Trp Pro Phe Phe
195 195 200 205
197 Ile Gly Ile Gly Thr Ala Ala Ile Gly Gly Ala Leu Ile Ala Leu Phe

```

## RAW SEQUENCE LISTING

DATE: 07/10/2000

PATENT APPLICATION: US/09/604,231

TIME: 15:45:58

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07102000\I604231.raw

```

198      210      215      220
200 Asn Ile Lys Ala Val Ala Leu Gly Ala Ala Gly Phe Leu Gly Val Val
201 225      230      235      240
203 Ser Ile Asp Ala Pro Asp Met Val Met Phe Leu Val Cys Ala Val Val
204      245      250      255
206 Thr Phe Phe Ile Ala Phe Gly Ala Ala Ile Ala Tyr Gly Leu Tyr Leu
207      260      265      270
209 Val Arg Arg Asn Gly Ser Ile Asp Pro Asp Ala Thr Ala Ala Pro Val
210      275      280      285
212 Pro Ala Gly Thr Thr Lys Ala Glu Ala Glu Ala Pro Ala Glu Phe Ser
213      290      295      300
215 Asn Asp Ser Thr Ile Ile Gln Ala Pro Leu Thr Gly Glu Ala Ile Ala
216 305      310      315      320
218 Leu Ser Ser Val Ser Asp Ala Met Phe Ala Ser Gly Lys Leu Gly Ser
219      325      330      335
221 Gly Val Ala Ile Val Pro Thr Lys Gly Gln Leu Val Ser Pro Val Ser
222      340      345      350
224 Gly Lys Ile Val Val Ala Phe Pro Ser Gly His Ala Phe Ala Val Arg
225      355      360      365
227 Thr Lys Ala Glu Asp Gly Ser Asn Val Asp Ile Leu Met His Ile Gly
228      370      375      380
230 Phe Asp Thr Val Asn Leu Asn Gly Thr His Phe Asn Pro Leu Lys Lys
231 385      390      395      400
233 Gln Gly Asp Glu Val Lys Ala Gly Glu Leu Cys Glu Phe Asp Ile
234      405      410      415
236 Asp Ala Ile Lys Ala Ala Gly Tyr Glu Val Thr Thr Pro Ile Val Val
237      420      425      430
239 Ser Asn Tyr Lys Lys Thr Gly Pro Val Asn Thr Tyr Gly Leu Gly Glu
240      435      440      445
242 Ile Glu Ala Gly Ala Asn Leu Leu Asn Val Ala Lys Lys Glu Ala Val
243      450      455      460
245 Pro Ala Thr Pro
246 465
249 <210> SEQ ID NO: 3
250 <211> LENGTH: 1109
251 <212> TYPE: DNA
252 <213> ORGANISM: Corynebacterium glutamicum
254 <220> FEATURE:
255 <221> NAME/KEY: CDS
256 <222> LOCATION: (1)..(1086)
257 <223> OTHER INFORMATION: FRXA00315
259 <400> SEQUENCE: 3
260 tat gat ttc ggc ggt cca gtc ggc ggt ctg ctc ttc ggt ctg gtc tac 48
261 Tyr Asp Phe Gly Gly Pro Val Gly Gly Leu Leu Phe Gly Leu Val Tyr
262 1 5 10 15
264 tca cca atc gtc atc act ggt ctg cac cag tcc ttc ccg cca att gag 96
265 Ser Pro Ile Val Ile Thr Gly Leu His Gln Ser Phe Pro Pro Ile Glu
266 20 25 30
268 ctg gag ctg ttt aac cag ggt gga tcc ttc atc ttc gca acg gca tct 144

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/604,231  
 DATE: 07/10/2000  
 TIME: 15:45:58

Input Set : A:\seqlist.txt  
 Output Set: N:\CRF3\07102000\I604231.raw

```

269 Leu Glu Leu Phe Asn Gln Gly Gly Ser Phe Ile Phe Ala Thr Ala Ser
270      35      40      45
272 atg gct aat atc gcc cag ggt gcg gca tgt ttg gca gtg ttc ttc ctg 192
273 Met Ala Asn Ile Ala Gln Gly Ala Ala Cys Leu Ala Val Phe Phe Leu
274      50      55      60
276 gcg aag agt gaa aag ctc aag ggc ctt gca ggt gct tca ggt gtc tcc 240
277 Ala Lys Ser Glu Lys Leu Lys Gly Leu Ala Gly Ala Ser Gly Val Ser
278      65      70      75      80
280 gct gtt ctt ggt att acg gag cct gcg atc ttc ggt gtg aac ctt cgc 288
281 Ala Val Leu Gly Ile Thr Glu Pro Ala Ile Phe Gly Val Asn Leu Arg
282      85      90      95
284 ctg cgc tgg ccg ttc ttc atc ggt atc ggt acc gca gct atc ggt ggc 336
285 Leu Arg Trp Pro Phe Phe Ile Gly Ile Gly Thr Ala Ala Ile Gly Gly
286      100      105      110
288 gct ttg att gca ctc ttt aat atc aag gca gtt gcg ttg ggc gct gca 384
289 Ala Leu Ile Ala Leu Phe Asn Ile Lys Ala Val Ala Leu Gly Ala Ala
290      115      120      125
292 ggt ttc ttg ggt gtt gtt tct att gat gct cca gat atg gtc atg ttc 432
293 Gly Phe Leu Gly Val Val Ser Ile Asp Ala Pro Asp Met Val Met Phe
294      130      135      140
296 ttg gtg tgt gca gtt gtt acc ttc ttc atc gca ttc ggc gca gcg att 480
297 Leu Val Cys Ala Val Val Thr Phe Phe Ile Ala Phe Gly Ala Ala Ile
298      145      150      155      160
300 gct tat ggc ctt tac ttg gtt cgc cgc aac ggc agc att gat cca gat 528
301 Ala Tyr Gly Leu Tyr Leu Val Arg Arg Asn Gly Ser Ile Asp Pro Asp
302      165      170      175
304 gca acc gct gct cca gtg cct gca gga acg acc aaa gcc gaa gca gaa 576
305 Ala Thr Ala Ala Pro Val Pro Ala Gly Thr Thr Lys Ala Glu Ala Glu
306      180      185      190
308 gca ccc gca gaa ttt tca aac gat tcc acc atc atc cag gca cct ttg 624
309 Ala Pro Ala Glu Phe Ser Asn Asp Ser Thr Ile Ile Gln Ala Pro Leu
310      195      200      205
312 acc ggt gaa gct att gca ctg agc agc gtc agc gat gcc atg ttt gcc 672
313 Thr Gly Glu Ala Ile Ala Leu Ser Ser Val Ser Asp Ala Met Phe Ala
314      210      215      220
316 agc gga aag ctt ggc tcg ggc gtt gcc atc gtc cca acc aag ggg cag 720
317 Ser Gly Lys Leu Gly Ser Gly Val Ala Ile Val Pro Thr Lys Gly Gln
318      225      230      235      240
320 tta gtt tct ccg gtg agt gga aag att gtg gtg gca ttc cca tct ggc 768
321 Leu Val Ser Pro Val Ser Gly Lys Ile Val Val Ala Phe Pro Ser Gly
322      245      250      255
324 cat gct ttc gca gtt cgc acc aag gct gag gat ggt tcc aat gtg gat 816
325 His Ala Phe Ala Val Arg Thr Lys Ala Glu Asp Gly Ser Asn Val Asp
326      260      265      270
328 atc ttg atg cac att ggt ttc gac aca gta aac ctc aac ggc acg cac 864
329 Ile Leu Met His Ile Gly Phe Asp Thr Val Asn Leu Asn Gly Thr His
330      275      280      285
332 ttt aac ccg ctg aag aag cag ggc gat gaa gtc aaa gca ggg gag ctg 912
333 Phe Asn Pro Leu Lys Lys Gln Gly Asp Glu Val Lys Ala Gly Glu Leu

```

## VERIFICATION SUMMARY

DATE: 07/10/2000

PATENT APPLICATION: US/09/604,231

TIME: 15:45:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07102000\I604231.raw

L:8 M:283 W: Missing Blank Line separator, <120> field identifier  
L:11 M:283 W: Missing Blank Line separator, <130> field identifier  
L:12 M:283 W: Missing Blank Line separator, <140> field identifier  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:14 M:283 W: Missing Blank Line separator, <160> field identifier  
L:1028 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 13, CDS LOCATION:77..631  
L:1492 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 19, CDS LOCATION:98..1606  
L:1894 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 25, CDS LOCATION:14..766